DOCUMENT RESUME

ED 343 960 TM 018 196

AUTHOR Shellman, David W.
TITLE One Paradox in District Accountability and Site-Based Management.
PUB DATE Feb 92
NOTE 10p.; Paper presented at the Annual Meeting of the North Carolina Association for Research in Education (Chapel Hill, NC, February 20, 1992).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Accountability; Communication Skills; Comparative Testing; Criterion Referenced Tests; Educational Assessment; Grade 6; Instructional Effectiveness; Intermediate Grades; Mathematics Achievement; *School Based Management; School Districts; Sex Differences; *Standardized Tests; State Programs; Testing Programs; *Test Use
IDENTIFIERS California Achievement Tests; *Instructional Management Systems; Mathematics and Communications Skills Prog NC; *North Carolina (Gaston County)

ABSTRACT

The paradox of site-based school management with use of standardized tests or instructional management systems that restrict teacher choices was evident in one school district in North Carolina in which measurement of student success has centered on student performance on state-mandated tests. A study was conducted to see if students whose teachers were part of the Gaston County (North Carolina) instructional management system for mathematics and communication skills (MAC)--a criterion referenced test-based system--scored higher on the California Achievement Test (CAT) than students whose teachers were not part of the MAC program. Subjects were 949 MAC sixth graders, with a control group of 1,306 non-MAC students. Data were collected from the 1988-89 administrations of the CAT. Students whose teachers used the MAC program did not outperform non-MAC students in any respect. In addition, male students in the MAC program consistently scored lower than did male students not in the MAC program and scored significantly lower than did female students in the MAC program. The attempt to improve test scores through the MAC program has not been successful. Recommendations are made for a complete evaluation of the MAC program. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED343960

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

DAVID W. SHELLMAN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

One Paradox in District Accountability and Site-Based Management

By

David W. Shellman

Presented At The Annual Conference

North Carolina Association for Research in Education

February 20, 1992

Chapel Hill, North Carolina

TM018196

BEST COPY AVAILABLE

One Paradox in District Accountability
And Site-Based Management
DAVID W. SHELLMAN
Gaston County Schools

David W. Shellman
Gaston County Schools
1608 Hoffman Road
Gastonia, NC 28054
Home: (704) 867-5879
Office: (704) 866-6225

Education administrators set the tone and establish quality workplaces for teachers and students. Additionally, administrators are ultimately responsible (accountable) for use of public funds and student success. In recent years, there has been public and policy-maker concern about student success and teacher "professionalization". Sometimes seeking these two ideals has produced unusual situations: site-based management but use of standardized tests or instructional management systems that restrict teacher choices. This experimental study addresses these paradoxes as they operated in one North Carolina (NC) school system.

Numerous reports have addressed the ills of education in America. A Nation at Risk by the National Commission on Excellence in Education and reports from the Carnegie Foundation were the "landmark beacons" around which restructuring that we now experience were designed. These plans, simple ones, call for restructuring of schools giving classroom teachers greater input into the decision making process while holding those same teachers far more accountable for these decisions and student progress than ever before. The concept has political overtones from the school house clear to the state house because coupled with this "new freedom" and accountability is the funding for public education.

Measurement of student success in NC has centered on student performance on state mandated tests. The purpose of this study was to determine if students whose teachers were a part of the Gaston County(NC) instructional management system for math and communication skills(MAC) -- a criterion-referenced test-based system -- scored higher on the *California Achievement Test*(CAT), a norm-referenced test than did students whose

teachers were not a part of the MAC Program. The study addressed the purpose by testing three hypotheses which were rejected or retained based upon an Alpha level of $p \leq 0.05$. The hypotheses are as follows: (1) there is no difference between the experimental (MAC) and control (Non-MAC) groups on CAT scores; (2) there is no difference between the experimental (MAC) and control (Non-MAC) groups on CAT scores as function of gender, race, or parental education level; (3) there is no difference between the experimental (MAC) and control (Non-MAC) groups on CAT scores as function of years of participation in the MAC Program.

The population of this study consisted of sixth grade students (N=2255) in the Gaston County public schools during the 1988-89 school year. The division of this sixth grade population into "control" and "experimental" groups began in 1986 with the implementation of MAC in a non-experimental setting. Schools with computers capable of supporting the state sponsored Student Information Management System (SIMS) and MAC were chosen to begin the MAC Program. The first year, six schools with sixth graders were in the MAC Program. The second year, 1987-88, six additional schools with sixth graders were added to the MAC Program. During the 1988-89 school year, no additional sixth graders were added to the program; the control group, non-MAC students (N=1306), represented approximately fifty-eight percent (58%) of the total sixth grade population; the experimental group of MAC students (N=949), represented approximately forty-two percent (42%) of the population. To accommodate a balanced design, 210 students from each subpopulation were selected randomly as participants in the study.

The data for this study were collected during the 1988-89 school year from tests results on the CAT purchased from CTB/McGraw-Hill. A Posttest Only Design was applied

to the data. In addition to the statistical procedures used to tests the hypotheses two additional procedures were used. To establish that the control and experimental groups were comparable a t-test was performed on students' score on the Test of Cognitive Skills, a group IQ test administered in North Carolina. The results of the t-test showed no difference between the mean IQ scores of the control and experimental groups. To estimate the internal consistency of the MAC math and communication skills tests a Kuder-Richardson Formula 20 was used on MAC data collected during the 1990-91 school year. Results estimated r to be 0.95 and 0.98, respectively.

For each of the hypotheses, all subtests of the CAT (vocabulary, comprehension, total reading, language mechanics, language expression, total language, math computation, math concepts, total math, spelling, and total battery) were used to determine if students in the experimental group outperformed students in the control group on any aspect of the CAT. Rejection of the null hypothesis would favor an alternate hypothesis that states there is a difference between the experimental and control groups on CAT scores.

Hypothesis One: In Reading, Language, and Spelling, t-test results show there are no differences between the control and experimental group means. In these areas students in the MAC Program do not outperform students not in the MAC Program. Assuming proper implementation of the MAC Program, one could then conclude that the MAC Program offers no beneficial feedback to teachers that improves instruction and thus improves CAT Reading, Language, and Spelling scores. For Math, a different condition exists. In all areas of Math, one rejects the null hypothesis in favor of the alternate hypothesis. In all cases, mean scores for the control-group students are significantly higher than means for the experimental group. MAC Program students scores significantly lower on Math than do non-MAC students. The

MAC Program, as it exists, does not offer the feedback necessary to improve Math scores on the CAT. The magnitude of the differential on Math scores caused the researcher to reject the null hypothesis for mean Total Battery scores.

Hypothesis Two: Gender: T-test results for female students show, there is no difference between the control and experimental group mean scores on Reading, Language, and Spelling. On Math Computation and Total Math, control group students outperformed the experimental (MAC) students. For the Total Battery score, one would retain the null hypothesis. For male students, one retains the null hypothesis on all Reading, Language Mechanics, and Spelling subtests. On Language Expression and Total Language, and all Math subtests one rejects the null hypothesis in favor of the control group. Finally, for males, one rejects the null hypothesis favoring the control group for the Total Battery score.

Race: On every subtest, one retains the null hypothesis for non-Caucasian students. For Caucasian students, one retains the null hypothesis on all Reading, Language, and Spelling subtests, but rejects it for every Math subtest and on Total Battery.

Parental Education Level: ANOVA results show for students whose parents have a high school education, one retains the null hypothesis for Reading and Spelling, and rejects the null hypothesis for Language, Math, and Total Battery in favor of the control group. MAC students whose parents have either less than a high school education or more than a high school education score as well as non-MAC students on the CAT. The only exception is on Math Computation where non-MAC students outperform MAC students significantly.

Hypothesis Three: Hypothesis three tests the longevity factor in the study. The control group has zero years in the MAC program while the experimental group is divided into two groups of students having two and three years of participation. ANOVA results show that

for Reading, Language Mechanics, and Spelling one retains the null hypothesis; and for Language Expression, all Math subtests, and Total Battery one rejects the null hypothesis favoring the control group.

Sign tests show that the overwhelming majority of the higher mean scores, although not all significant, were higher for Non-MAC than for MAC. All significantly different scores favored the Non-MAC means.

Findings/Conclusions

1. The MAC Program does not help Gaston County students to score higher on the CAT in the sixth grade. On no test did students whose teachers used the MAC Program outperform students whose teachers did not use the MAC Program. Although the literature supports measurement-driven instruction as a measure of student achievement, this study's findings are in contrast to the literature. The MAC Program is not a positive factor on student test scores on the CAT. The findings support this conclusion. In cases where significant differences were found, all favored students who were not in the MAC Program.

2. This study indicates that a differential impact on sixth-grade male students exits. Experimental-group males scored significantly lower than control-group males and experimental-group females on CAT subtests. One expects that an instrument used to measure student progress to be free of bias against any particular group. Male students in the MAC Program consistently scored lower than male students not in the MAC Program and significantly lower than female students in the MAC Program. No significant differences between males and females in the control group were detected on any CAT subtests.

3. The MAC Communication Skills and MAC Math Tests are not educationally sound instruments to measure student progress. Although both MAC tests have an extremely high reliability estimate (math: $r = 0.94$; communications: $r = 0.98$), both tests violate the concepts of good criterion-referenced tests. These violations include number of items for determination of mastery, number of objectives tested at one time, and (possibly) validity.

4. The attempt to improve tests scores through the MAC Program is a failure for Gaston County Schools. Gaston County has invested more than $1,000,000 and seven years in the MAC Program with very few positive results. Each year teachers take between two and two and one-half weeks of instructional time to administer the MAC pre- and posttests.

5. The MAC Program with its measurement-driven instruction concept takes too many decisions from the classroom teacher in a site-based management environment. The new reform in education calls for more decision making at the classroom level that affects student outcomes. Measurement-driven instruction is far too structured for site-based management because it makes most of the instructional decisions from the teacher through its prescriptive, diagnostic testing.

Discussion

Gaston County has invested approximately $1,000,000 and countless hours in the development, implementation, and maintenance of the MAC Program. For a commitment of this proportion in an educationally sound program, one would expect that Gaston County students should benefit from this project and show improvement in test scores. This clearly

is not the case in this study. Results of this study show that, at best, students in the MAC Program do equally as well on the CAT as students not in the MAC Program.

Teacher influence on student outcomes is difficult to address and was not attempted in this study. An assumption of this paper was that teachers assigned to the MAC Program were using test results as tools for planning instruction. Students in the MAC Program did not out perform students not in the program. One could conclude that MAC is a deterrent to the education process, but to be sure, the researcher would need to add a study component to assess how and to what degree the teachers actually use the MAC Program.

Another factor not controlled by this study was the historical path of teachers a student had prior to the sixth grade. Although t-test on IQ scores established the groups to be of similar aptitude no attempt to control teachers or quality of instruction was made. Finally there was a group of teachers in Gaston County that had no access to an instructional management system (MAC) yet the students they taught outperformed students who had the benefit of MAC.

Recommendations For Action

This study investigates three pertinent hypotheses related to the Gaston County Instructional Management System known as the MAC Program and student results obtained on the CAT. No attempt has been made to explain why the results obtained in this study present such a paradoxical result as they do. In light of the enormous investment by Gaston County Schools of both its fiscal and personnel resources one would expect the experimental group to outperform the control group. This clearly is not happening. At this point the researcher recommends the following to Gaston County:

1. A complete formative and summative evaluation of the MAC Program should be conducted by an independent consultant to answer such questions as cost-benefit, teacher use of MAC, teacher satisfaction, strengths and weaknesses of the program, test/retest reliability, and then to recommend continuation or termination of the MAC Program.
2. Policy should be developed and enacted by the Board of Education that outlines procedures for developing, field testing, piloting, implementing, and evaluating new instructional programs in Gaston County thus, hopefully, preventing problems such as the MAC to be inflicted on the students of Gaston County.
3. Based only on the results of this study the researcher recommends that Gaston County discontinue use of the MAC Program if no further study shows evidence for continuing MAC.